Personalized Ontological Framework for Web Information Retrieval

Smita R.Sankhe.
University of Mumbai
K.J.Somaiya College of Engineering

Kavita Kelkar.
University of Mumbai
K.J.Somaiya College of Engineering

ABSTRACT
In proposed system personalized ontology for web information retrieval is introduced: Specificity and Exhaustively. Specificity describes a subject's focus on a given keyword. Exhaustively restricts a subject's semantic space dealing with the topic. Personalized ontology framework is proposed for knowledge representation and reasoning over behavior of users. This framework learns user profiles from both a world knowledge base and user background knowledge. The world knowledge and user background information are used to attempt to discover and specify user background knowledge. From a world knowledge base (WordNet database) personalized ontology are constructed focusing on user occupation. Ontological framework provides a solution to emphasizing global and local knowledge in a single computational framework. We present a personalized user specific ontological framework using WordNet knowledge for web information retrieval which will help to present the relevant search result to the user.

General Terms

WordNet, local knowledge, Global Knowledge, web information retrieval.

Keywords

Ontology, personalization, semantic relations, world knowledge, Background knowledge, user profiles, web information gathering.

1. INTRODUCTION
As a model for knowledge description and formalization, ontologies are widely used to represent user profiles in personalized web information retrieval. However, when representing user profiles, many models have utilized only knowledge from either a global knowledge base or user local information. In this proposed project, a personalized ontology framework is proposed for knowledge representation and reasoning over behavior (occupation) of users. This framwork learns ontological user profiles from both a world knowledge base and user background knowledge. The amount of web-based information available has increased dramatically. How to gather useful information from the web has become a challenging issue for users. Current web information retrieval systems attempt to satisfy user requirements by capturing their information needs. For this purpose, user profiles are created for user background knowledge description .User profiles represent the concept models possessed by users for retrieval of web information. A concept framework is implicitly possessed by users and is generated from their background knowledge. While this concept model cannot be proven in laboratories, many web ontology have observed it in user behavior. When users read through a document, they can easily determine whether or not it is of their interest or relevance to them, a judgment that arises from their implicit concept models. If a user's concept model can be simulated, then a superior representation of user profiles can be built. To simulate user concept models, ontologies—a knowledge description and formalization model—are utilized in personalized web information retrieval. Such ontologies are called ontological user profiles or personalized ontologies. To represent user profiles, many researchers have attempted to discover user background knowledge through global or local analysis. Global analysis uses existing global knowledge bases for user background knowledge representation. Commonly used knowledge bases include generic ontologies (e.g., WordNet), online knowledge bases (e.g., online categorizations and Wikipedia). The global analysis techniques produce effective Performance for user background knowledge extraction. However, global analysis is limited by the quality of the used knowledge base. For example, WordNet was reported as helpful in capturing user interest in some areas but useless for others[3]. Local analysis investigates user local information or observes user behavior in user profiles. In some works, such as, users were provided with a set of links and asked for relevance feedback. User background knowledge was then discovered from this feedback for user profiles. However, because local analysis techniques rely on data mining or classification techniques for knowledge discovery, occasionally the discovered results contain noisy and uncertain information. As a result, local analysis suffers from ineffectiveness at capturing formal user knowledge [1][2]. From this it can be hypothesized that user background Knowledge can be better discovered and represented if we can integrate global and local analysis within a hybrid framework.

2. DISCUSSION
This section discusses various existing model for personalized ontological web information retrieval.

2.1 Golden Model: TREC Model

The TREC model [1] was used to demonstrate the interviewing user profiles, which reflected user concept models perfectly. For each topic, TREC users were given a set of documents to read and judged each as relevant or nonrelevant to the topic. The TREC user profiles perfectly reflected the users' personal interests, as the relevant judgments were provided by the same people who created the topics as well, following the fact that only users know their interests and preferences perfectly.

2.2 Baseline Model: Category Model

This model demonstrated the non-interviewing user profiles, a user's interests and preferences are described by a set of weighted subjects learned from the user's browsing history.

These subjects are specified with the semantic relations of super class and subclass in ontology. When an OBIWAN agent receives the search results for a given topic, [1][2]it filters and reranks the results based on their semantic similarity with the subjects. The similar documents are awarded and re-ranked higher on the result list.

2.3 Baseline Model: Web Model

The web model [1]was the implementation of typical semi interviewing user profiles. It acquired user profiles from the web by employing a web search engine. The feature terms referred to the interesting concepts of the topic. The noisy terms referred to the paradoxical or ambiguous concepts.

3. PROPOSED SYSTEM

Proposed project work is provisioned to following phases. This process exactly gives idea about how it will design and generate the ontology framework for web information retrieval. The proposed ontology framework aims to discover user background knowledge and learns personalized ontology using WordNet to represent relevant search result. Fig. 1 illustrates the architecture of the ontological framework. A personalized ontology is constructed, according to a given topic. Two knowledge resources, the global world knowledge base (wordNet Database) and the user's background knowledge, are utilized by the framework. The world knowledge base provides the meaningful classes for personalized ontology. The user background knowledge (occupation) is discovered from the user profile.

Personalized ontology's that formally describe and specifies user background knowledge. For example a user searching for a word might have different expectations, for searching the same query. For example if we are searching for the term "New Jersey", business travelers may expect different search from leisure travelers. A user may become a business traveler when planning for a business trip, or a leisure traveler when planning for a family holiday. A user's concept model may change according to different information needs.

3.1 Global Knowledge Representation

World Knowledge representation research involves analysis of how to accurately and effectively reason and how best to use a set of symbols to represent a set of facts within a knowledge domain. In this model user background knowledge is extracted from a world knowledge base encoded from the Word Net.First, step is the construction of world knowledge base. The world knowledgebase must cover the wide range of topics, since Users expect different results for searching a single word query. The Word Net was developed for Organizing and retrieving information from a large volume of library collections

3.2 Methodology

Working of the proposed system is presented as follows in six steps:

Step 1: User should be registered after registration local information of that particular user will get stored in the local database (local information of the user) to identify behavior of the user this will help to display relevant result to the user as per his/her occupation.

Step 2: After registration user profile will get generated and that information will help to exact meaning full keyword from global knowledge base.

Step 3: This step is main crux of system where actually key word generation using WordNet (ontology learning) is take place for exacting meaningful classes from the WordNet database.

Step3.1: **Keyword Generation**: By using wordNet data base administrator has been generate the keywords for searching the particular information.

Step3.2: **Keyword Knowledge**: Before searching the keywords administrator must be stored the correct information for particular key word with help of ontology process. This keyword information has been stored in wordNet database.

Step 3.3: **Mapping of Local and Global Database**: in this step local information of user (occupation) will get mapped with all the keyword extracted from global knowledge base (WordNet) with semantic relation and then Administrator has been provide the url for particular search key word with this url will get result from local database.

Step 4: **Search Keyword**: Authenticated user will be search the key word in this framework. For example users want particular meaning of the keyword that result will be providing the ontology process.

Step 5: **Relevant Search Result**: Based on the user occupation they will get result of the particular keyword. For example student wants only meaning of the particular key word and employee wants the some information about particular keyword so based on the requirement will get the results

Step 6: **Feedback (Ranking):** Once user get the result, to check whether user is satisfied with the result or not user

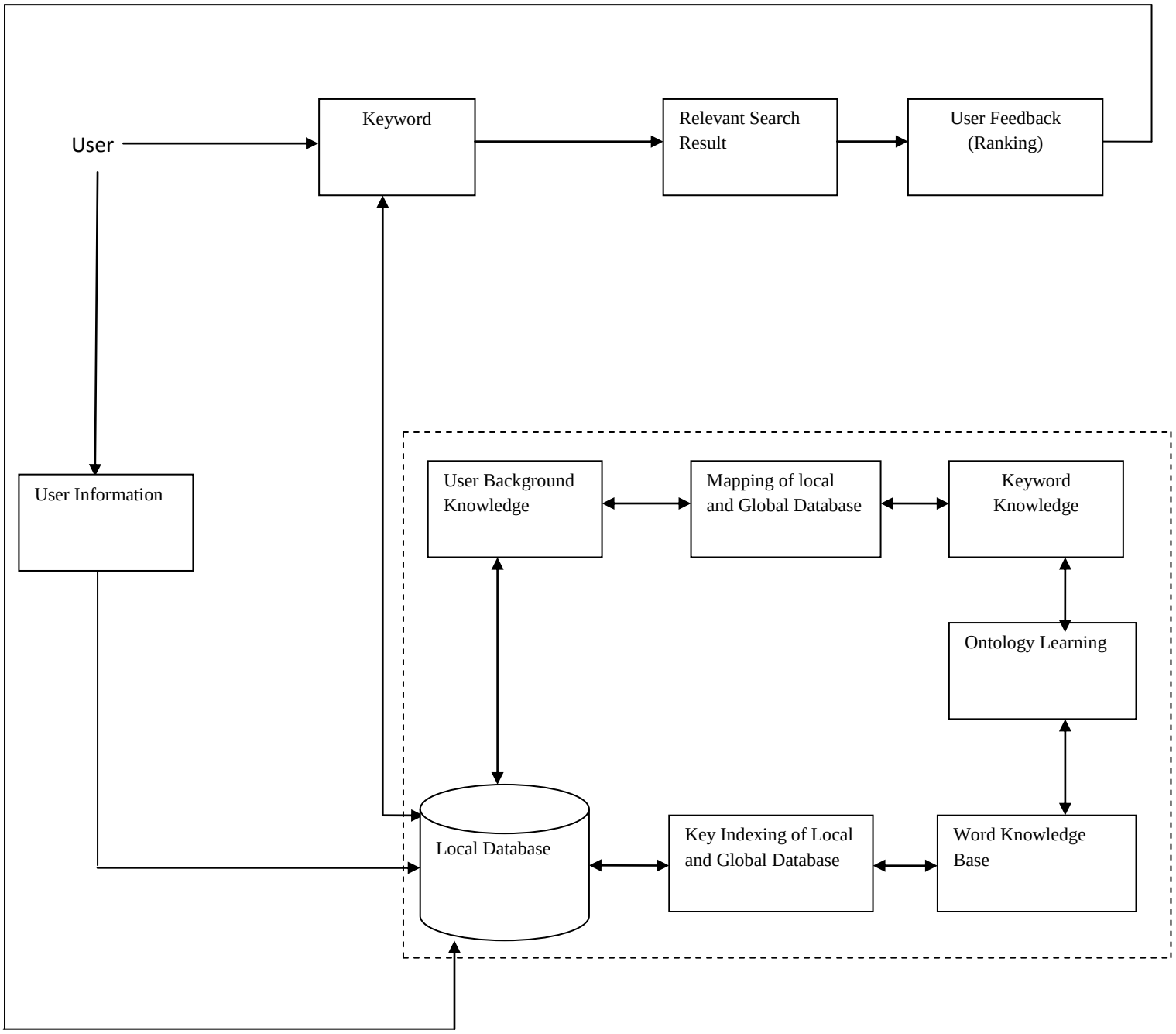

Fig 1: Proposed System Architecture

3.3 Design of ontology Using WordNet

The world knowledge and a user's background knowledge employed in the projected framework. World data is commonsensical data acquired by folks from expertise and education and local instance repository may be a user's personal assortment of data things. From a world cognitive content, this tend to construct customized ontology's using WordNet by adopting user behaviour.

concepts are inter-related which collectively impose a structure on the domain and constrain the possible interpretations of terms.

Ontology may take a variety of forms, but necessarily it will include a vocabulary of terms, and some specification of their meaning. This includes definitions and an indication of how

WordNet-based ontology is to select a manageable number of classes that have sufficient conceptual depth to enable effective semantic inference and enough variety to yield the widest lexical coverage. This work have carried out to date is primarily concerned with verbs, but the approach developed extends to other word classes in WordNet (nouns, adjectives and adverbs) in a straightforward manner. In defining an event ontology based on WordNet, we selected verb synonym sets

that were less specific in meaning as event classes (e.g., {communicate#2, intercommunicate#2} vs. {gesticulate#1, gesture#1, motion#1}). In doing so, we chose the more frequent member of the synonym set to name the class, e.g. communicate#2 for the synonym set {communicate#2, intercommunicate#2}. [3][11]

3.4 Comparative Study

Table 1. Comparison of Existing system with proposed system

Existing System	Proposed System
Search using either local or global knowledge base	Search using both local and global knowledge base
Quires are based on OWL model	Quires are based upon WordNet model
Representation is done through Documents	Representation is done through Textual Information with links

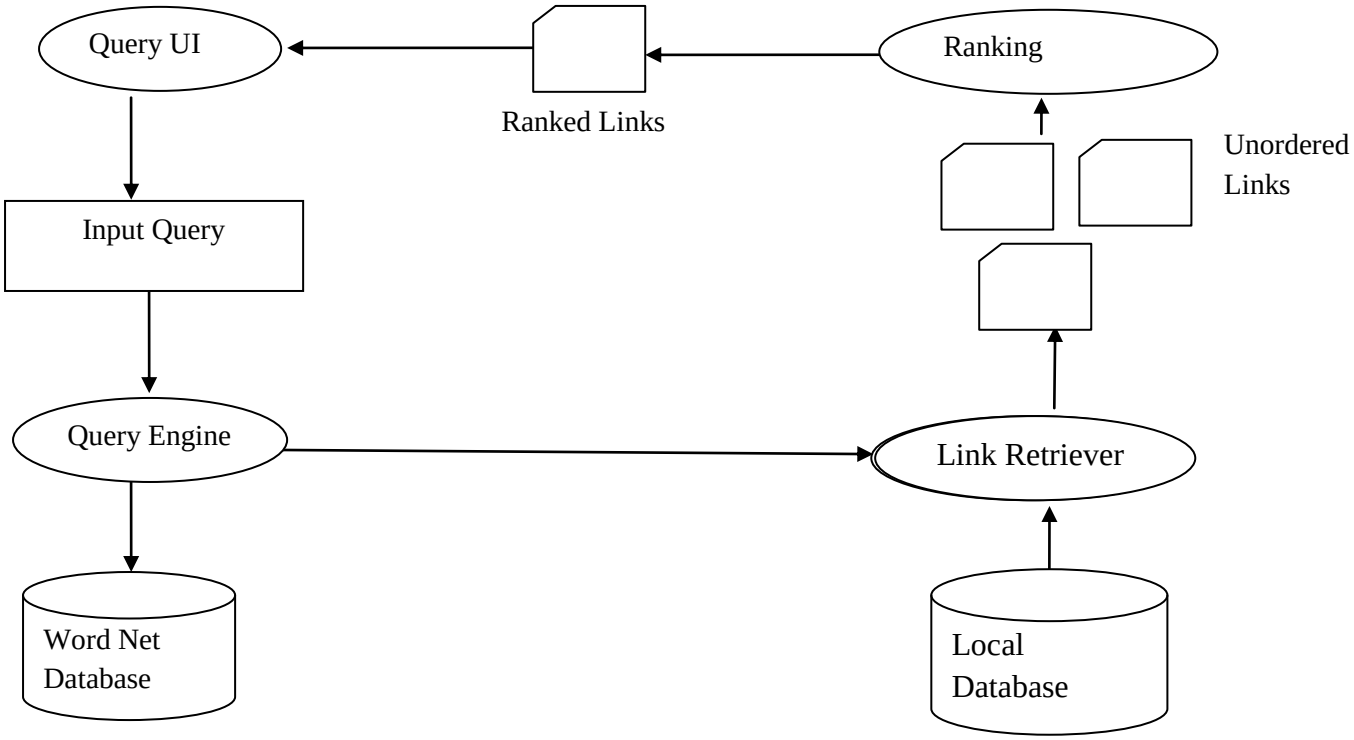

Fig 2: Design Process of Ontology using WordNet Database

The above fig2. As represented by the design process of the ontology web information retrieval .Our system takes as input a query. This query could be generated from a keyword query, as in a natural language query, a form-based interface where the user can explicitly select ontology classes and enter property values, or more sophisticated search interfaces. A number of research works have undertaken the construction of easy to use user interfaces for ontology query languages, and we do not address this problem here. The input query is executed against the knowledge base, which returns a list of instance tuples that satisfy the query. Finally, the documents that are annotated with these instances are retrieved, ranked, and presented to the user.

4. CONCLUSION

Result of the proposed system is search result in form of url that will help to user for web information retrieval. Based on the user occupation they will get result of the particular keyword. For example student wants to only meaning of the particular key word and employee wants the some information about particular keyword so based on the requirement will get the results. To generate this result mapping of global

knowledge base (WORDNET DATABASE) [12]and user lrticular key word and employee wants the some information about particular keyword so based on the requirement will get the results. To generate this result mapping of global knowledge base (WORDNET DATABASE) [11] and user local information (occupation) is required

5. REFERENCES

[1] Xiaohui Tao,Yuefeng Li,and Ning Zhong,"A personalized Ontology Model for web information Gathering June 2011.

[2] Maria Golemati, Akrivi Katifori,and Costas Vassilakis, , Creating an Ontology for the User Profile: Method and Applications, In Proceedings of the First International Conference on Research Challenges in Information Science, Dec 2007.

[3] Jaap Kamps,"Visualizing WordNet Structure" July 2005.

[4] L.M. Chan, Library of Congress Subject Headings: Principle and Application. Libraries Unlimited, May 2005.

[5] John D.King "Mining word knowledge for analysis of search engine content",pp 7-15,April 2005.

[6] P.A. Chirita, C.S. Firan, and W. Nejdl, "Personalized Query Expansion for the Web," Proc. ACM SIGIR ('07), pp. 7-14, May 2007.

[7] A. Doan, J. Madhavan, and A. Halevy, "Learning to Map between Ontologies on the Semantic Web," Proc. 11th Int'l Conf. World Wide Web (WWW '02), pp. 662-673, June 2002.

[8] D. Downey, S. Dumais, and E. Horvitz, "Understanding the Relationship between Searchers' Queries and Information Goals," Proc. 17th ACM Conf. Information and Knowledge Management (CIKM '08), pp. 449-458, December 2008.

[9] E. Frank , G.W. Paynter, "Predicting Library of Congress Classifications from Library of Congress Subject Headings," J. Am. Soc. Information Science and Technology, vol. 55, no. 3, pp. 214-227,Augest 2004.

[10] S. Gauch, J. Chaffee, and A. Pretschner, "Ontology-Based Personalized Search and Browsing," Web Intelligence and Agent Systems, vol. 1, nos. 3/4, pp. 219-234, June 2003

[11] J.D. King, Y. Li, X. Tao, and R. Nayak, "Mining World Knowledge for Analysis of Search Engine Content," Web Intelligence and Agent Systems, vol. 5, no. 3, pp. 233-253, May 2007.

.